IAF ASTRODYNAMICS SYMPOSIUM (C1)
Mission Design, Operations & Optimization (1) (8)

Author: Mr. Haiyang Li
Tsinghua University School of Aerospace, China, lihy15@mails.tsinghua.edu.cn

Ms. Zhemin Chi
Tsinghua University, China, chizhemin1021@163.com
Prof. Hexi Baoyin
Tsinghua University, China, baoyin@tsinghua.edu.cn

MULTI-RENDEZVOUS TRAJECTORY OPTIMIZATION WITH NEURAL NETWORK AND REINFORCEMENT LEARNING

Abstract

This paper presents a deep neural network and reinforcement learning method to optimize multi-rendezvous spacecraft trajectory. The crucial part of multi-rendezvous trajectory optimization is finding the ideal sequence, which is a complex combinatorial optimization problem, and one of the examples is the multiple debris removal problem of 9th Global Trajectory Optimization Competition released by ESA. The threat of space debris has been more and more concerned recently, thus multi-rendezvous debris removal optimization is considered in this paper. The chemical propulsion system is utilized, and the single transfer is implemented by multi-impulse. This optimization problem can be divided into two subproblems. The first subproblem is to estimate the fuel consumption because solving each multi-impulsive transfer accurately in large-scale fast searches for the optimal sequence is not practical. The second subproblem is to find the optimal sequence with a powerful global optimization algorithm. For the first subproblem, a deep neural network (DNN) is trained to estimate the fuel consumption. The database is generated by solving 10000 multi-impulsive transfers considering J2 perturbation. The input of the DNN is initial equinoctial elements, final equinoctial elements, and transfer time. The output of the DNN is fuel consumption of this transfer. The DNN consists of 2 hidden layers of 256 units each with ReLu activations. After well-tuned, the DNN can estimate the fuel consumption with very little error. For the second subproblem, a recurrent neural network (RNN) is trained using a model-free policy gradient method. The reinforcement learning paradigm is followed to tackle this combinatorial optimization problem, using the performance index of removal missions as the reward signal. Compared to supervised learning, reinforcement learning is preferable because optimal labels of such problems are not accessible but the quality of a set of solutions can be compared. 302 pieces of Iridium-33 debris objects are considered in this paper, and the original data of debris clouds can be found from the celestrak. Two kinds of missions are optimized. The first is to maximize the removal benefit gained by one single spacecraft, and the second is to minimize the number of spacecraft needed to clear up a group of debris, for example, a group of 67 pieces of debris which are large in size. Simulations demonstrates that the proposed method performs better than other algorithms.